

明 細 書

質量分析方法および質量分析装置

技術分野

本発明は質量分析装置を用いて、特にタンパク質の同定精度の向上や処理時間の短縮を図るための装置および方法に関するものである。

背景技術

タンパク質やペプチドを同定する場合、エドマン分解法や質量分析法を用いた方法が一般的に知られている。

質量分析装置を用いてタンパク質の同定を行うものの例として特表平 9-510780 号公報がある。この例は、タンパク質やペプチドのアミノ酸配列情報を集積した一般に公開されているデータベースからマスペクトルを予測し、さらに、実際に測定を行って得られた試料のマスペクトルと比較を行い、一致度を元に同定を行うことが開示されている。

また、質量分析装置のデータ処理において、測定結果を一般のデータベースを用いて解析する例として、特開平 5-164751 号公報がある。

発明の開示

従来の方法においては、次に挙げる問題があった。

(a) 未知試料におけるタンパク質の同定精度

測定によって得られたイオンの質量には、タンパク質のアミノ酸配列を集積したデータベースにおいて、多くのペプチドが対応している可能

性があった。このような場合、そのペプチドを含むタンパク質の数が多いため、タンパク質をそのペプチドから特定することは困難である。

(b) 同定処理の時間

タンパク質やペプチドの同定は、測定後のデータ処理として実施される。もし、タンパク質やペプチドのアミノ酸配列情報のデータベースを利用する場合、データベースの内容の変化によって、同定精度は直接影響を受ける。例えば、測定データは同じでも、最新のデータベースを適用することによって、それまで未知であったペプチドやタンパク質が同定できる場合がある。そのような場合を想定すると、データベースが更新されるごとに同定処理を行う必要がある。データベースの登録内容が激増し、測定データも着実に増える状況においては、同定に要する処理時間は飛躍的に増大してしまう。したがって、同定処理に要する時間を短縮できる手法が強く望まれている。

(c) 同定に必要な試料量

生体の細胞内に存在するタンパク質は多量に存在するものから、ごく微量のものまで様々である。もし、ごく微量に存在するタンパク質の同定が目的であれば、測定が可能な量を確保しなければならない。しかし、生物種や組織によっては入手困難かつ高価なものが多い。

また、質量分析装置を用いてタンパク質の同定を行うためには、液体クロマトグラフ質量分析装置 (LC/MS) やマトリクス支援レーザー脱離イオン化質量分析装置 (MALDI-MS) を用いる場合が多いが、LC/MS では、LC から溶出された試料が連続的に MS へ導入され、測定されるマススペクトルが時々刻々と変化する。MALDI-MS では、予め前処理した試料をマイクロプレート等に配置してレーザーを照射しイオン化するため、試料が微量である場合には、出現するイオンも極

僅かであり、マススペクトルとして検出できる時間も非常に限られている。

質量分析装置を用いてタンパク質の同定を行う場合は、より詳細なマススペクトルを得るために、一度取得したイオンを開裂させて得る所謂、MS/MSスペクトルを一般に用いる。得られるマススペクトルが常に変化する場合、或いは非常に極僅かの時間しか出現しない場合は、全てのイオンにおいて一度にMS/MSスペクトルを得ることが出来ない場合がある。特に、目的のタンパク質やペプチドが極微量しかない場合は、何度も繰り返し試料を導入してマススペクトルを得ることは不可能である。従って、より極僅かな試料しかない状態でもMS/MSスペクトルを用いて同定処理が可能な手法が望まれている。

(d) 特定のタンパク質の選択性

質量分析装置によってタンパク質やペプチドを測定する場合、特定のタンパク質に着目して同定したい場合がある。例えば、タンパク質の機能や構造上の特性、細胞内での局在情報、発現パターン、特定の疾患との関連などによって特定のタンパク質を選択することができる。無作為にタンパク質やペプチドを同定するのではなく、それらのタンパク質の選択性を高め、それらを同定する確率を向上させることは重要である。

また、特定のタンパク質以外の選択性を高めることができれば、結果的に夾雑物や翻訳後修飾されたペプチドの同定確率の向上に結びつくと考えられる。

本発明の目的は、質量分析装置を用いてタンパク質やペプチドの同定を行う場合に、より高速且つ高確度に同定を行うことが可能な質量分析方法および装置を提供することである。

上記目的を達成するための本発明の特徴は、試料をイオン化し、質量

分析装置を用いてタンパク質の分析を行う質量分析方法において、タンパク質やペプチドの情報が格納されたデータベースより、所定の情報を選出し、当該選出された成分の質量を推測し、質量ごとの頻度情報を算出し、試料を質量分析装置によって分析してマススペクトルを得、当該得られたマススペクトルと前記頻度情報に基づいて同定に用いるべき質量を選出し、当該選出された質量のマススペクトルを前駆イオンとして質量分析し、得られたマススペクトルを用いて同定処理を行うことである。

図面の簡単な説明

第1図は、第1の実施例の処理フローを示す図である。

第2図は、第1の実施例において集積した頻度情報を示すグラフである。

第3図は、本発明の装置構成例である。

第4図は、条件設定画面の例である。

第5図は、前駆イオンの選択結果表示画面例である。

第6図は、第2の実施例の処理フローを示す図である。

第7図は、第2の実施例において集積した頻度情報を示すグラフである。

第8図は、第3の実施例の処理フローを示す図である。

第9図は、第2の実施例において集積した頻度情報を示すグラフである。

発明を実施するための最良の形態

以下、本発明の実施例について説明する。

◎第1の実施例

○装置構成

第3図に、本発明が適用される質量分析計およびデータ処理装置の構

成を示す。

本実施例の装置構成は、試料の分離のためのクロマトグラフ装置 10、質量分析装置本体 11、その制御部 15、およびデータ処理部 17を、信号線 16 で結んだものである。質量分析装置本体 11 は、試料をイオン化するためのイオン源 12、質量分析部 13、および検出部 14 から構成されている。また、データ処理部 17 には、キーボード 18 と、表示装置 19 が具備されている。

データ処理装置 17 は、インターネット等の外部の公衆回線に接続可能であり、これにより、ネット上に接続されたデータベースにアクセスし、必要な情報を入手することが出来る。また、データベースの情報は、CD-ROM等の記録メディアを用いて入手しても良い。

またこの実施例では、クロマトグラフ装置とイオントラップ型の質量分析計を図示したが、質量分析計には、MS¹分析によって前駆イオンを選択しMS²分析を行うこと（所謂、MS/MS分析）が可能な全ての質量分析計が適用可能である。イオン源には、タンパク質やペプチドをなるべく壊さないようにイオン化できるイオン源が好ましく、例えばエレクトロスプレーイオン源（ESI: ElectroSpray Ionization）を用いることが出来る。また、クロマトグラフ装置は必ずしも必要ではなく、マトリクス支援レーザー脱離イオン化法（MALDI: Matrix Assisted Laser Desorption Ionization）などを用いる質量分析計であっても応用できる。

○処理内容

第1図において本実施例の基本的な処理を示す。尚、本実施例におけるタンパク質の同定には、タンパク質を所定の消化酵素で酵素消化して生成したペプチドを質量分析し、得られたマスペクトルの各ピークを

更に開裂し質量分析して（MS／MS分析）得られるマスペクトルを用いるものとする。第1図では、MS／MS分析のための前駆イオンの選択手法について示している。

また、ここでは酵素消化されたヒト由来タンパク質を試料として想定している。

以下、第1図の流れに沿って説明する。

（a）タンパク質のデータベース

まず、比較対照のタンパク質のデータベースとして、アミノ酸配列が登録されているものを選定する。なお、ヒト由来のタンパク質に着目するのであれば、少なくとも生物種に関する情報が含まれているデータベースを選定する必要がある。

一例として、米国NCBI（National Center for Biotechnology Information）が公開している、nr（Non-redundant protein database）などが該当する。これは、各種のタンパク質のデータベースから、タンパク質のアミノ酸配列を収集したものである。

（b）特定タンパク質のアミノ酸配列集積

上記（a）のデータベースから測定の目的に合うタンパク質のアミノ酸配列を集積する。

例えば、NCBIのnrに登録されているものから、スイスSIB（Swiss Institute of Bioinformatics）およびEBI（The European Bioinformatics Institute）が作成したタンパク質のデータベース（Swiss-Prot）由来のものを抽出する。さらにその中から、生物種をヒトに限定しアミノ酸配列を集積する。

（c）酵素消化後のペプチドのアミノ酸配列導出

上記（b）によって集積したタンパク質を所定の消化酵素によって酵

素消化された後に生成されるペプチドのアミノ酸配列を求める。求めた配列はタンパク質ごとに整理する。なお、適用する消化酵素の種類は、事前に指定する。

(d) ペプチドの分子イオンの質量計算

上記(c)で求めた各タンパク質のペプチドが質量分析計において分子イオンとして観測される際の質量を計算する。

なお、質量の値は少数点以下を四捨五入し、整数化したものを想定する。また、個々の質量に対するペプチドが1個存在する場合1、存在しない場合0を与える。同一の質量を有するペプチドが複数ある場合、その質量の縦軸の値には一致するペプチドの個数を与える。

(e) 質量と頻度集積

上記(d)で求めた各タンパク質のデータを質量ごとに集積し、質量と頻度の関係として集積する。なお、以下の説明では、整数質量を m 、頻度を F とし、 m に対応する F の値を F_m と表す。

第2図は、2002年12月に取得したnrからSwiss-protのヒト由来配列を抽出し、さらにトリプシンで消化することを想定して求めた質量ごとの頻度である。この図でも明らかなように、質量ごとの頻度の値には大きなばらつきが認められる。すなわち、質量によって該当するペプチドの数が大きく異なることを示している。

(f) 質量ごとの重み計算

上記(e)で求めた質量ごとの頻度重みのパターンを計算する。ここでは次の二つのパターンを想定している。

整数質量(m)ごとの重みを W_m とすると、その値は次のようにして求めることができる。

$$W_m = 1 / (F_m + 1) \quad \dots (1)$$

この（１）式では、頻度が０に近いほど重みが大きくなる。この計算結果を重みパターン１とする。

続いて、次の式によって重みパターン２を求める。

$$W_m = 1 - 1 / (F_m + 1) \quad \dots (2)$$

この場合、重みパターン１とは反対に、頻度の値が大きいほど重みも大きくなる。

このように、本実施例では、２種類の重みパターンを作成している。それぞれは、頻度が小さいものを強調するパターンと、頻度が大きいものを強調するものである。

これまでの処理は、質量分析計による測定に先立ち、事前に準備しておく。（ｇ）以降は、実際の試料に対して行うものである。

（ｇ）MS¹分析

ヒト由来の試料を上記（ｃ）と同一条件で酵素消化し、MS¹分析を行う。

（ｈ）MS¹スペクトル

MS¹分析によって得られるMS¹スペクトルは、一般に質量対電荷比と強度の関係として求めることができる。ここで観測されるイオンが前駆イオンの候補となる。

（ｉ）質量抽出

上記（ｈ）で得られたMS¹スペクトルから、質量の情報を抽出する。具体的には、各質量においてMS¹スペクトルにあらかじめ指定した閾値以上のピークが存在すれば１、存在しなければ０を与える。すなわち、該当する質量に対応するピークの有無を０，１の２値で示す。なお、ここでは重みパターンと整合性を保つため、整数化した質量を採用する。２値化された値をEとすると、整数質量（m）ごとのEの値はE_mと表

することができる。

(j) 擬似スペクトル作成

上記 (i) で得たピークの有無と、(f) で求めた重みパターンから、擬似スペクトルを生成する。重みパターンを W_m 、ピークの有無を E_m とすると、擬似スペクトルの強度 I_m は次式で求めることができる。

$$I_m = W_m \times E_m \quad \dots (3)$$

この式では、擬似ピークの強度は重みパターンの値を反映したものとなる。また、2種類の重みパターンを採用する場合、擬似スペクトルも2種類作成する。すなわち、重みパターン1は擬似スペクトル1に、重みパターン2は擬似スペクトル2に反映される。

(k) 前駆イオン選択とMS/MS分析

MS²分析を行うべき前駆イオンを、(j) で求めた擬似スペクトルの強度に着目して選択する。ここでは、擬似スペクトル1, 2および両方を利用し前駆イオンを選択する例を方式1, 2, 3に分けて説明する。

方式1では、擬似スペクトル1を利用して前駆イオンを選択する。すなわち、擬似スペクトル1の強度に着目し、その値の大きい順に前駆イオンとして採用する。ここで、擬似スペクトル1は低頻度の質量を強調した重みパターン1が採用されている。すなわち、ここで選択した前駆イオンの質量は、上記(c)で求めた酵素消化後のペプチドのアミノ酸配列において低頻度に出現するものである。

第1図に示した例では、③→⑤→①→④→②の順に前駆イオンの質量が決定され、MS/MS分析が実施される。

また、方式1においては、頻度が0の場合、擬似スペクトルの強度が1となる。これは、データベース内に対応するペプチドの情報が存在しないことを意味する。測定後のタンパク質の同定処理を想定した場合、

最低でも一つのペプチドが対応していることが望ましいので、強度 1 となる質量は除外すべきである。また、頻度が大きい場合、擬似スペクトルの強度は 0 に近い値になる。この場合は、低頻度の質量に着目する意味が薄れてしまう。このように、前駆イオンの選択においては、強度の範囲が重要となる。そこで、事前に擬似スペクトルの強度の範囲を指定しておき、その範囲において大きいものから選択することが必要である。

もし、擬似スペクトルの強度が 1（頻度が 0）となる質量を前駆イオンとして選択した場合、上記（c）で質量を求めた計算条件に該当しない未知のタンパク質由来のペプチドである可能性がある。また、該当する場合でも、翻訳後修飾により質量が増減している場合なども考えられる。逆に、そのようなものを優先して取得したい場合には、前駆イオンの対象を擬似スペクトルの強度が 1 のものに限定する。

また、擬似スペクトルの強度が 1（頻度が 0）となる質量を除外した上で擬似スペクトルの強度が高いものから前駆イオンを選択した場合、該当するペプチドの情報がデータベース中に存在し、その数が比較的少ないことを示している。これは、測定後の後処理においてペプチドやタンパク質を同定する際に、少ない候補から絞り込むことができることを意味する。すなわち、ペプチドやタンパク質の同定精度の向上に寄与する。目的の成分を同定するまで測定を繰り返すような場合には、試料の量の低減に直接寄与することができる。さらに、候補となるペプチドが少ないことより、絞り込みに要する時間の低減にも寄与できる。

方式 2 は、高頻度の質量を強調した擬似スペクトル 2 から前駆イオンを選択するものである。この場合も、強度の大きなものから順に前駆イオンを選択する。この方式によって前駆イオンを選択した場合、その質量は上記（c）で求めた酵素消化後のペプチドのアミノ酸配列において

高頻度に出現する。すなわち、測定後のタンパク質の同定において、該当するタンパク質を幅広く選択したい場合に向いている。

なお、第1図の例では、②→④→①→⑤→③の順番で前駆イオンが選択され、MS/MS分析が実施される。

方式2の場合、同じ配列情報を有するタンパク質群として、異なる生物種において共通する祖先を持つオルソログ遺伝子由来のタンパク質、種内で遺伝子重複後に変異が蓄積されたパラログ遺伝子由来タンパク質、およびゲノム上の同じ遺伝子領域から異なるmRNAが生成されるスプライス・バリエント由来のタンパク質などについての情報を得る可能性が高まる。

方式3は、方式1と方式2のそれぞれで順序付けられた前駆イオンを交互に採用するものである。第1図の例では、方式1の選択結果を優先し、③→②→⑤→④→①の順番を得ている。この場合は、低頻度と高頻度の質量を順に指定することにより、測定後のタンパク質の同定において、タンパク質の同定精度を上げるとともに、幅広くタンパク質を選択したい場合などに向いている。

上記の方式1～3の方式のどれを採用するか、また方式3において高頻度を優先するか低頻度を優先するかなどの設定は、測定に先立ち指定しておく。尚、上記の説明においてはアイソトープ（同位体）並びに多価イオンの影響を省略している。

また、本実施例では、実試料を質量分析して得られたマススペクトル中の全てのピークを前駆イオンとするのではなく、順位付けした質量の先頭から何番目の質量までを前駆イオンとするのかについて予め設定する機能を備える。LC/MSで分析を行う場合やMALDI-MSなどで極微量の試料を分析する際は、①～⑤のような前駆イオンの候補とな

るピークが観測される時間は限られているため、全てのマススペクトルのピークについてMS / MS分析を行うための時間が取れない場合が多い。このような場合に本実施例のように有効と思われる質量から順序付けを行い、前駆イオンとする手法は非常に有効である。

すなわち、MALDI-MSなどの場合には前駆イオンの候補となるピークが観測されるであろう時間を勘案し、前駆イオンとして選択する数として、その時間中に測定可能な範囲内でなるべく大きい値を設定する。LC / MSなどの場合、前駆イオンの候補が観測される時間内に分析を終える必要があることから、MS分析ならびにMS / MS分析に要する時間を考慮した上で、前駆イオンとして選択する数を設定する。

(1) MS / MSスペクトルを用いた同定処理

最後に、上記(k)のステップで選択した前駆イオンのMS / MSスペクトルを用いて同定処理を行う。同定処理に関しては、上記(c)のステップで集積したデータと、前駆イオンを分析して得られたMS / MSスペクトルとを用いて行う。

以上が一連の処理の流れである。

ここで、第3図に示した装置構成と、第1図に示す処理内容との対応を次に示す。第1図における特定タンパク質のアミノ酸配列集積から質量ごとの重み計算までをデータ処理部17において実施する。ここで作成された重みパターン1と2を、試料の分析が開始される前に制御部15に転送しておく。測定に際しては、制御部15において試料のMS¹分析が行われ、MS¹スペクトルが取得される。そこで、質量抽出、擬似スペクトル作成、および前駆イオン選択の各処理を行う。制御部15では、選択された前駆イオンに対しMS²分析を行い、MS²スペクトルを取得する。場合によっては、さらに次の前駆イオンからMS²スペ

クトルを取得する工程を繰り返す。以上が、第1図の手法を第3図の装置構成に適用した場合の基本的な流れとなる。

以上のように、本実施例においては、1回の試料導入によって MS^1 分析と MS^2 分析を行うことを前提としている。もし、同じ試料を複数回に分けて導入できる場合、1回目の測定において全ての MS^1 スペクトルを求めた後で、一括して本手法による前駆イオンの選択を行い、2回目以降の測定を行うこともできる。その場合は、前駆イオンの選択を制御部15で行うのではなく、データ処理部17が担当する。

また、測定に際しては、前駆イオンの指定を測定開始からの時間などに依存して切り替える手法が適用できる。

○測定条件設定画面

試料を測定する前に実施する各種条件の設定画面を第4図に示す。なお、基本的には第1図に示す方法に則り、第2図のデータ処理部17における表示装置19およびキーボード18によって実現することを想定している。

第4図に示す画面において、タンパク質選択条件、頻度および重みパターンの計算条件と結果、および前駆イオン選択および MS/MS 分析条件の設定等を可能としている。以下に、それぞれの概要を示す。

(a) タンパク質選択条件

この部分では、タンパク質のデータベースから特定のタンパク質のアミノ酸配列を集積するための条件として、データベースおよび生物種をリストから選択し、目的の機能などに相当するキーワードを設定する。この例では、データベースとして“Swiss-prot”を、生物種として“homo sapiens”（ヒト）を選択している。また、キーワードには、“zinc finger”を入力している。なお、Zinc Finger とは、DNAに結合するタンパク

質に見られる機能部位である。そのような機能部位を有するタンパク質は、DNAからmRNAへの転写機能に関わっている可能性があると考えられている。すなわち、Swiss-prot のヒト由来のタンパク質から、zinc finger に関するものを集積するよう指定したものである。

(b) 頻度および重みパターン計算条件

この部分では、上記 (a) によって指定された特定のタンパク質のアミノ酸配列を集積した後、質量に対応する頻度情報を計算するための条件を設定する。まず、対応する修飾の種類、および消化酵素を選択する。続いて、分子イオンの質量に影響を及ぼすイオン化方法を選ぶ。さらに頻度を計算するための質量精度と、計算範囲、および質量の定義を指定している。

なお、修飾の設定においては詳細を省略しているが、その部位や修飾される確率などの指定も重要となる。さらに、消化酵素についても、消化が不十分な場合についての指定などが考えられるが、ここでは省略した。

(c) 頻度および重みパターン計算結果

この部分では、頻度、重みパターン 1、重みパターン 2 の中から、指定された計算結果をグラフ表示する。なお、グラフは、必要に応じて拡大、縮小できることが望ましい。

(d) 前駆イオン選択およびMS/MS分析条件

ここでは、前駆イオンを選択するための各種条件を設定する。

まず、前駆イオンの質量を限定するための質量範囲の入力、およびMS¹のマスペクトルにおけるイオン強度の閾値を設定する。続いて、擬似スペクトルから前駆イオンを選択する際の強度範囲を、低頻度と高頻度の場合について指定する。さらに、前駆イオンの選択条件を、“低

頻度から”，“高頻度から”，“高→低を交互に”，“低→高を交互に”の中から選択する。また、前駆イオンとして選択すべき数を回数や時間などから指定できるよう、MS／MS分析の繰り返し条件として、数値と単位の指定を可能とした。単位はプルダウンメニューから指定できる。回数を指定した場合、数値はMS／MS分析を行う回数を意味する。即ち、「前駆イオン」の項目で指定された方式で順序付けられたイオンの内、優先度の高いものから何番目までのイオンを前駆イオンとするかの指定である。時間を指定した場合には、MS分析によって前駆イオンを選択してから、最後のMS／MS分析が終わるまでの時間を示している。すなわち、ここで設定された時間を超えた場合、次の候補のイオンはMS／MS分析を行わないものとする。

○測定結果の表示

第5図に前駆イオン選択結果のグラフとテーブルを示す。

グラフには予めデータベースから抽出した情報から得られる頻度情報と、実際に試料を分析して得られるマススペクトルをオーバーラップさせて表示する。また、前駆イオンとして選択したイオンはピークトップに●印を、非選択のものは◆印を付け区別した。このグラフも必要に応じて拡大、縮小できることが望ましい。

テーブル表示においては、マススペクトルの質量、イオン強度、頻度、および選択の可否を示している。ここでは、第4図で前駆イオンの選択が“低頻度から”と設定されていることに合わせて、頻度の低いものから順にソートして表示し、頻度に着目しやすくしたテーブルを示す。

この表示画面によれば、測定後にデータ処理部17において、マススペクトルから選択した前駆イオンと前駆イオンの候補、および適用した頻度情報が表示装置19に同時に表示されるため、指定した条件で前駆

イオンが選択されたかどうかを検証することができる。

◎第2の実施例

本実施例は、あらかじめ集積したタンパク質のアミノ酸配列から、それぞれのタンパク質の質量と頻度として整理するものである。

この実施例も、基本的には第1図に示す手法と同様の流れになる。ただし、集積された特定タンパク質のアミノ酸配列から質量と頻度として集積するまでの部分が異なる。（第1図と異なり、酵素消化後のペプチドのアミノ酸配列を算出しない。）第6図にその部分を示す。

第6図の基本的な流れに沿って説明する。

（a）特定タンパク質のアミノ酸配列

第1図に示す特定タンパク質のアミノ酸配列と同様に、タンパク質のデータベースから生物種などの指定によって目的のタンパク質に対応するアミノ酸配列を集積する。

（b）タンパク質の分子イオンの質量計算

上記（a）で集積した各タンパク質が分子イオンとして観測される際の質量を計算する。なお、質量の値は少数点以下を四捨五入し、整数化したものを想定する。個々の質量に対する分子イオンが存在する場合1、存在しない場合0を与える。

（c）質量と頻度集積

上記（b）で求めた各タンパク質のデータを集積し、質量ごとの頻度として集積する。

第7図は Swiss-prot のヒト由来タンパク質から、さらに“Zinc Finger”と記載されたものを集積し、頻度情報として整理したものである。すなわち、この図は集積されたタンパク質の分子量分布を示している。

タンパク質は、分子量が非常に大きいものが多いため、一般的な四重極質量分析計やイオントラップ型の質量分析計では測定が困難であり、このような装置を用いる場合は、第1の実施例のようにペプチドに酵素消化したものを前提に同定を行う。これに対して本実施例においては、タンパク質そのものを測定できる質量分析装置（例えば、飛行時間型質量分析計（TOF-MS））を用いることを想定したものであり、このような装置を用いて本実施例のような頻度情報を適用することによって、ペプチドに酵素消化せずに、特定の質量のタンパク質を選択的に分析することが可能となる。

◎第3の実施例

本実施例では、特定のタンパク質に対する選択性を向上させる例を示す。特に、ここでは、遺伝子の発現に関わると考えられている“Zinc Finger”と称する機能部位を有するタンパク質に着目した。なお、この実施例は、第1図の実施例を拡張したものである。

第8図に概要を示す。この例では第1図における質量と頻度集積までの流れを多重化し、それらの結果から頻度差を求めている。また、第9図に頻度の計算例としてグラフを挙げた。

以下、第8図の流れに従い概要を説明する。

（a）タンパク質のデータベース

第1図に示す特定タンパク質のアミノ酸配列と同様に、タンパク質のデータベースから生物種などの指定によって目的のタンパク質に対応するアミノ酸配列を集積する。

（b）タンパク質の選択条件1による処理

上記（a）のタンパク質のデータベースから、選択条件1によって特定のタンパク質を選択し、アミノ酸配列の集積、酵素消化後のペプチド

のアミノ酸配列の導出，ペプチドの分子イオンの質量計算，質量と頻度の集積の各処理を行う。この流れは、第1図に示すそれぞれの処理と同じものである。

第9図のグラフ（A）は、タンパク質の選択条件1による処理結果の例である。これは、Swiss-prot のヒト由来タンパク質を集積し、トリプシンで消化した場合の頻度を示している。

（c）タンパク質の選択条件2による処理

上記（a）のタンパク質のデータベースから、選択条件2によって特定のタンパク質を選択し、アミノ酸配列の集積，酵素消化後のペプチドのアミノ酸配列の導出，ペプチドの分子イオンの質量計算，質量と頻度の集積の各処理を行う。この流れは、第1図に示すそれぞれの処理と同じものである。

上記（b）のタンパク質の選択条件1による処理とは、タンパク質を選択する際の条件が異なっている。第9図のグラフ（B）は、タンパク質の選択条件2による処理結果の例である。これは、Swiss-prot のヒト由来タンパク質から、さらに“Zinc Finger”の記載があるものを集積し、それをトリプシンで消化した場合を想定した。

（d）頻度差の計算

上記（2）および（3）で求めた質量と頻度の関係を質量ごとに減算し、その差を求める。なお、双方の頻度の総数が異なる場合、その比を乗じて是正するなどの処理を行う。

頻度差の計算例として、第9図のグラフ（C）を示す。これはグラフ（B）からグラフ（A）を質量ごとに減算したものである。ただし、グラフ（A）と（B）では、ペプチドの総数が異なるため、ここではグラフ（B）の側にそれぞれのグラフのペプチド総数の比を乗じて調整した。

このグラフ（C）において、例えば667や1072などの質量が突出している。これらの質量には、多くの Zinc Finger に共通するペプチドが対応している可能性がある。ちなみに、1072に対応するペプチド（IHTGEKPYK；アミノ酸の1文字表記）は Zinc Finger によく見られるものであり、このペプチドを前駆イオンとして選択することは、Zinc Finger の同定において非常に重要となる。

すなわち、グラフ（C）の頻度情報を活用し、頻度の数値の大きい質量を前駆イオンとして選択した場合、Zinc Finger 関連のタンパク質の選択性を向上させることができる。

この例では、データベースに登録されているタンパク質のアミノ酸配列をそのまま適用している。しかし、タンパク質のアミノ酸配列には、高頻度に繰り返されるものが知られている。従って、場合によっては、このような繰り返し配列を無視するなどの処置が望ましい。

その他の応用として、生物種、タンパク質の機能、構造上の特性、細胞内での局在部位、発現パターン、特定の疾患、データの出処などを考慮し、タンパク質の選択条件を決定することにより、それらのタンパク質由来のペプチドの選択性を向上できる可能性がある。

また、シグナル伝達などの研究分野においては、チロシンのリン酸化が重要な意味を持つ。そこで、着目するタンパク質において、チロシンがリン酸化されたペプチド断片に対する選択性を向上させる手法は有用と考える。非修飾のタンパク質において頻度情報を作成した場合、修飾によって質量差が生じるため、低頻度の質量を選択することによって修飾されたペプチドの選択性が高まる。さらに、厳密に非修飾と修飾について頻度情報の差を求める手法も有意である可能性がある。

生体組織に微量に存在するタンパク質を精製することは困難である。

特に、夾雑物となるタンパク質やペプチドを同定する必要がある場合、目的のもの以外の選択性を高めることは有用となる。この場合も、目的のタンパク質について頻度情報を作成し、低頻度の質量に着目して前駆イオンを選択することにより、夾雑物の選択性が向上すると考えられる。

以上に示す通り、本発明は、分析の目的に合致した頻度情報を用いて同定処理に必要なMS/MSスペクトルを得るための前駆イオンを効率良く選択することができ、同定精度の向上や処理時間の短縮を図ることが可能となる。

請 求 の 範 囲

1. 試料をイオン化し、質量分析装置を用いてタンパク質の分析を行う質量分析方法において、

タンパク質やペプチドの情報が格納されたデータベースより、所定の情報を選出し、当該選出された成分の質量を推測し、質量ごとの頻度情報を算出し、

試料を質量分析装置によって分析してマススペクトルを得、当該得られたマススペクトルと前記頻度情報に基づいて同定に用いるべき質量を選出し、当該選出された質量のマススペクトルを前駆イオンとして質量分析し、得られたマススペクトルを用いて同定処理を行うことを特徴とする質量分析方法。

2. 請求項1において、

前記データベースから入手した成分情報の質量を推定する際に、各タンパク質を所定の消化酵素によって酵素消化した後のペプチドを推定し、ペプチド毎に質量を推定することを特徴とする質量分析方法。

3. 請求項1において、

前記前駆イオンとして選択する数を予め設定することを特徴とする質量分析方法。

4. 試料をイオン化し、質量分析装置を用いてタンパク質の分析を行う質量分析方法において、

(A) タンパク質に関する情報を格納した外部のデータベースから複数のタンパク質の情報を入手するステップと、

(B) 上記入手したタンパク質ごとに質量を推定するステップと、

(C) 前記推定した質量が存在すれば1、存在しなければ0として正規化し、全てのタンパク質の推定結果を合計して質量ごとの頻度とし

て求め、重みパターンを算出するステップと、

(D) 試料を測定し、質量スペクトルを得るステップと、

(E) 試料のスペクトルに対し、質量が存在すれば1、存在しなければ0として正規化するステップと、

(F) 前記重みパターンに対して前記正規化した実試料のスペクトルを重ねし、擬似スペクトルを作成するステップと、

(G) 前記作成された擬似スペクトルに基づき、前記試料のマススペクトルの中からMS/MS分析すべき前駆イオンを選択するステップとを有することを特徴とする質量分析方法。

5. 請求項4において、

前記重みパターンを算出する際に、前記頻度が高い順に重み付けを行う第1のパターンと、前記頻度の低い順に重み付けを行う第2のパターンを作成することを特徴とする質量分析方法。

6. 請求項5において、

前記第2のパターンにおいて、重み付けの値が最高値を示す質量は、前記MS/MS分析すべき前駆イオンを選択する際に、除外されることを特徴とする質量分析方法。

7. 請求項4において、

上記データベースから情報を入手する際の条件を複数設定し、各条件設定によって得られたそれぞれの入手情報に対して上記(B)および(C)のステップを行い、各条件設定の重みパターンの頻度における差分を求め、新たな重みパターンを得ることを特徴とする質量分析方法。

8. 試料をイオン化するイオン化部、質量分析を行う質量分析部、分析条件の設定や分析結果のデータ処理を行うデータ処理部を備えた質量分析装置において、

前記データ処理部は、

予め設定された条件に従って、タンパク質に関する情報を格納したデータベースから情報を入手し、入手したタンパク質に対してピークが存在する質量の数をカウントして頻度情報として纏める事前準備処理と、

実試料を質量分析して得たマススペクトルに対応させ、前記頻度情報で得た頻度に応じてMS/MS分析すべき前駆イオンを選択する前駆イオン選択処理を行うことを特徴とする質量分析装置。

9. 請求項8において、

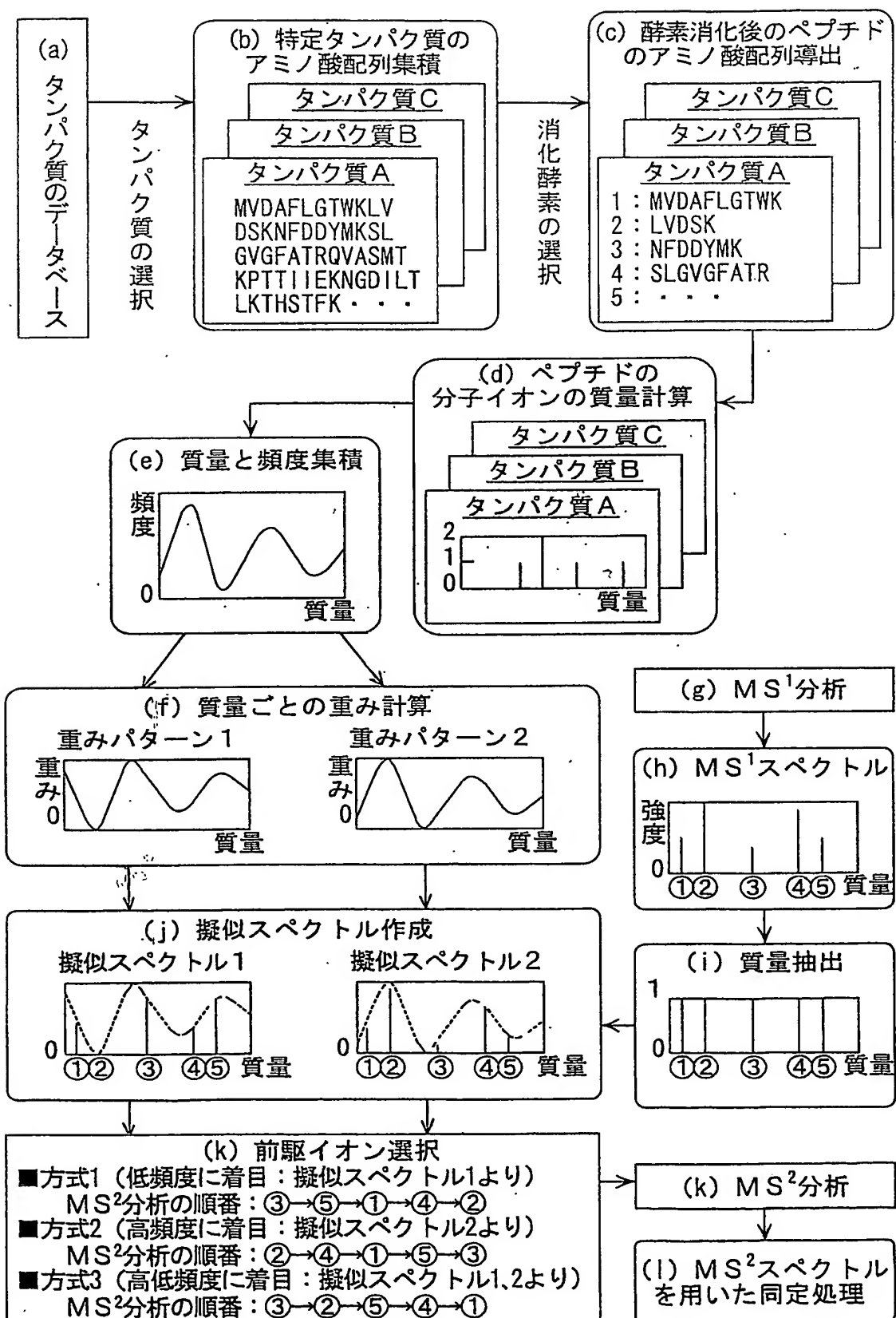
前記データ処理部は、表示部を有し、

当該表示部に、前記質量ごとの頻度情報を表示し、且つ当該表示内容に前記実試料を質量分析して得られたマススペクトルを重畳表示することを特徴とする質量分析装置。

10. 請求項8において、

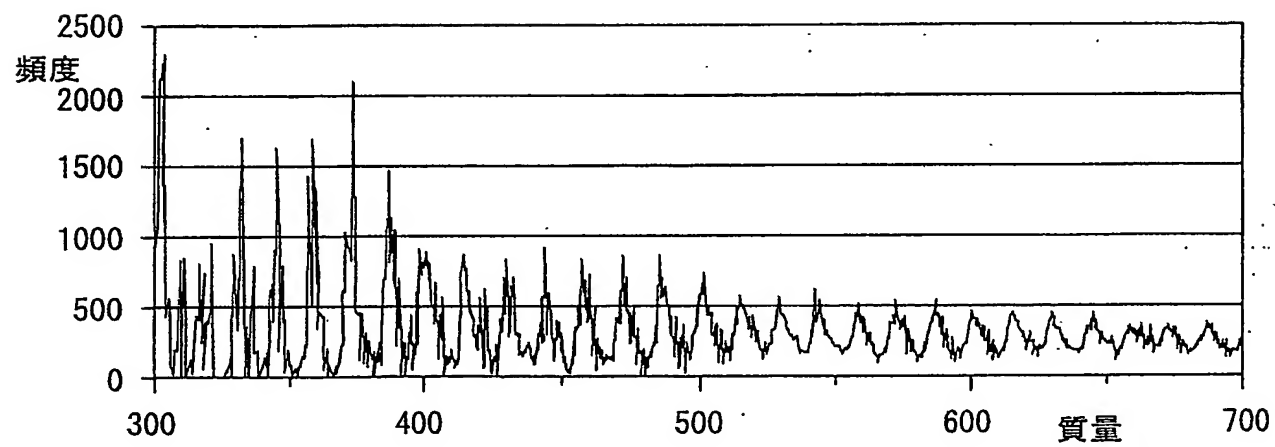
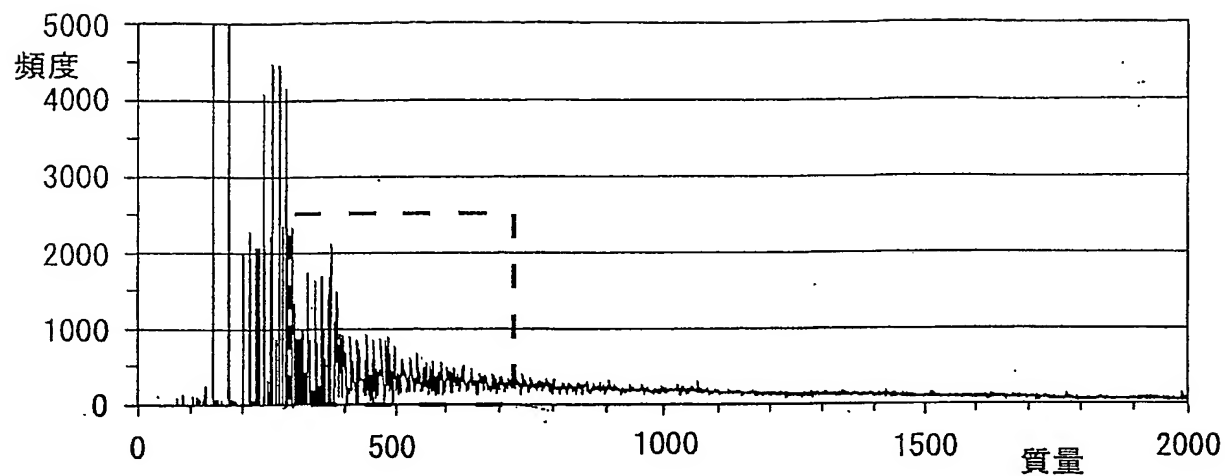
前記実試料を質量分析して得られたマススペクトル中の各ピークに対して、前駆イオンとして選択されたか否かを表示することを特徴とする質量分析装置。

第1図



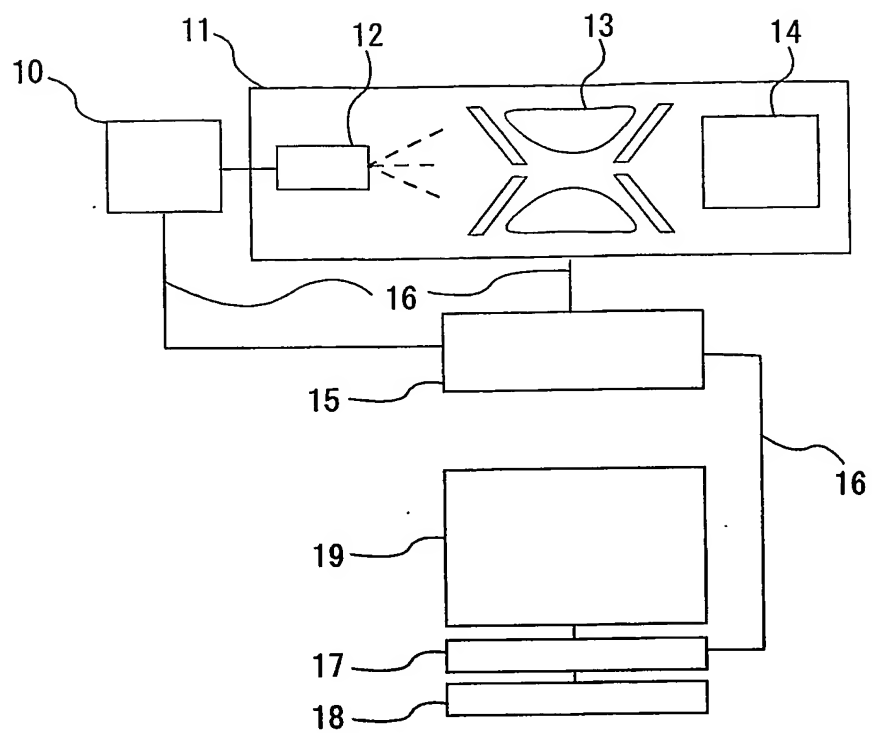
2 / 9

第2図



3 / 9

第3図



4 / 9

第4図

(a) タンパク質選択条件

データベース ▼ 生物種 ▼

キーワード

(b) 頻度および重みパターン計算条件

修飾 ☐ リン化 ☐ アセチル化 ☐ メチル化

消化酵素 ▼

イオン化方法 ▼

質量精度 (Da)

質量範囲 (Da) —

質量タイプ ☒ モノアイソトピック ☒ アベレージ

(c) 頻度および重みパターン計算結果

☒ 頻度 ☒ 重みパターン1 ☒ 重みパターン2

(d) 前駆イオン選択およびMS/MS分析条件

質量範囲 (Da) —

イオン強度の閾値 % ▼

擬似スペクトル強度 (低頻度) —

擬似スペクトル強度 (高頻度) —

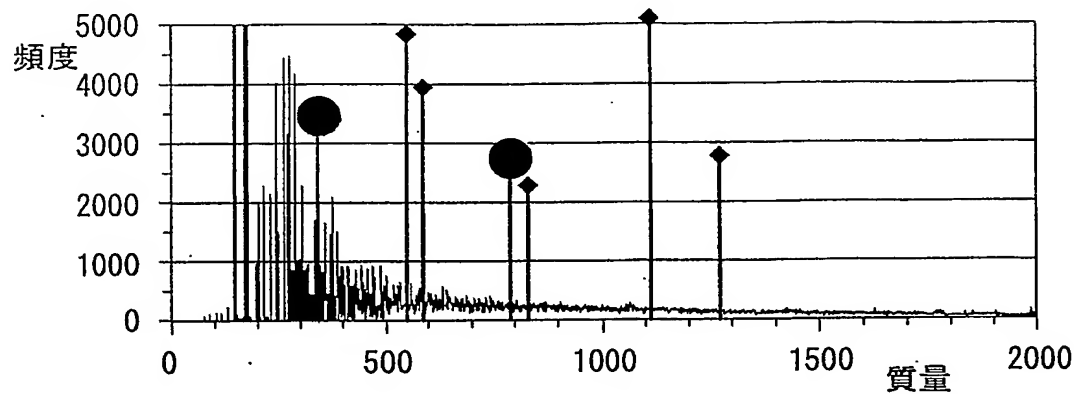
前駆イオン ☒ 低頻度から ☒ 高頻度から
☒ 高→低を交互に ☒ 低→高を交互に

MS/MS分析繰返し 回 ▼

5 / 9

第5図

前駆イオン選択結果グラフ

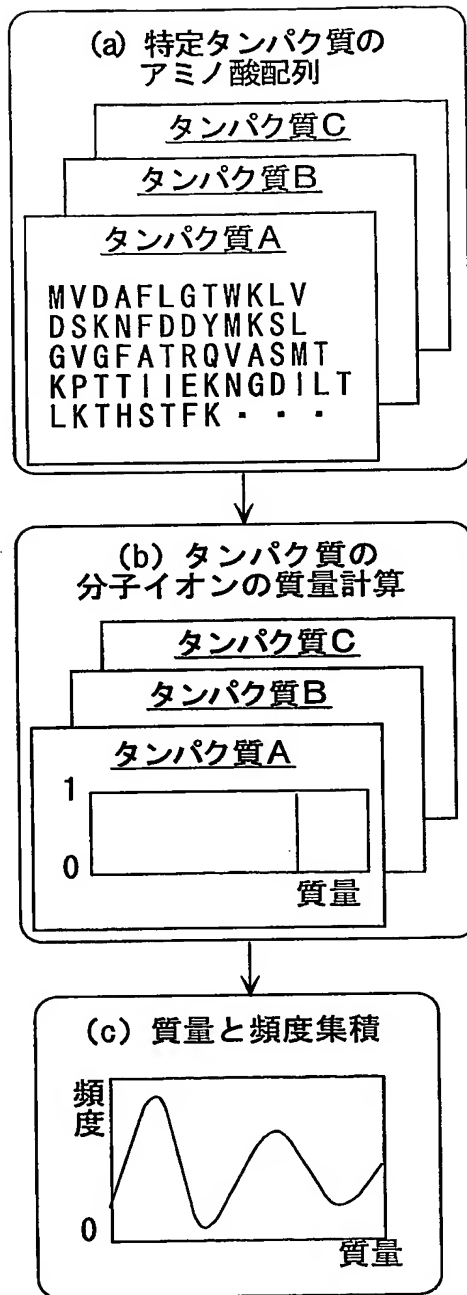


前駆イオン選択結果テーブル

No.	質量	イオン強度 (%)	頻度	選択
1	395	70	203	Yes
2	790	55	257	Yes
3	1241	55	260	No
4	1092	100	265	No
5	837	43	271	No
6	604	73	429	No
7	572	92	674	No

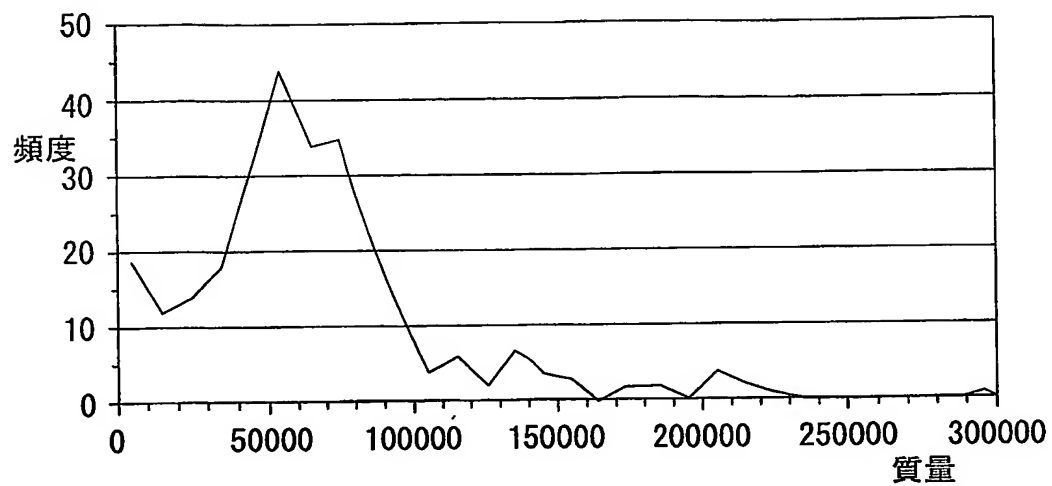
6 / 9

第6図



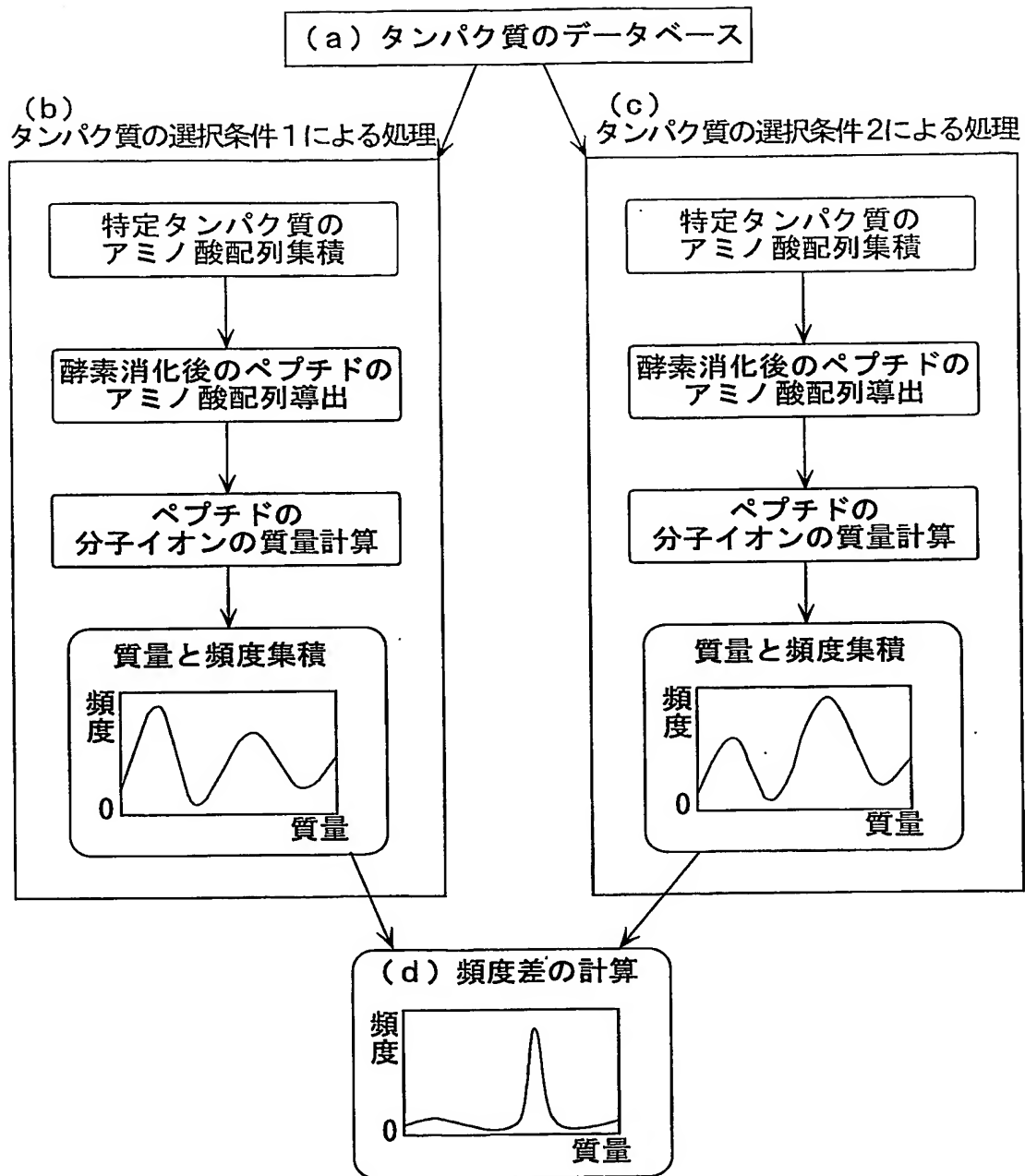
7 / 9

第7図



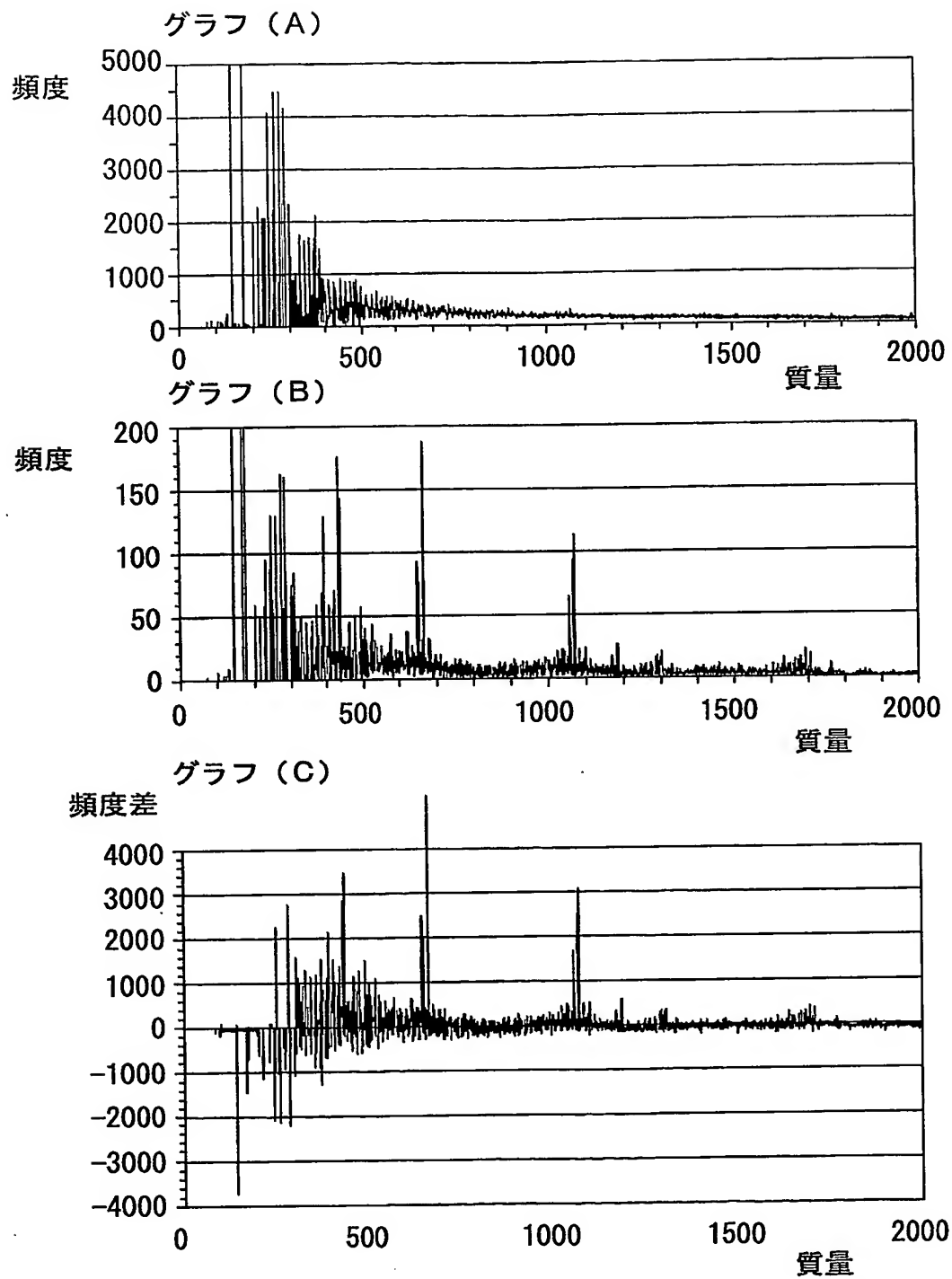
8 / 9

第 8 図



9 / 9

第9図



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/07923

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G01N27/62, H01J49/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G01N27/62-27/70, H01J49/00-49/48, G01N33/48-33/98,
C12Q1/00-1/70

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2003
Kokai Jitsuyo Shinan Koho	1971-2003	Jitsuyo Shinan Toroku Koho	1996-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICST FILE (JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95/25281 A1 (UNIVERSITY OF WASHINGTON), 21 September, 1995 (21.09.95), Full text; Figs. 1 to 6 & JP 9-510780 A	1-10
A	JP 8-124519 A (Shimadzu Corp.), 17 May, 1996 (17.05.96), Full text; Figs. 1 to 3 (Family: none)	1-10
A	JP 5-164751 A (Shimadzu Corp.), 29 June, 1993 (29.06.93), Full text; Figs. 1 to 3 (Family: none)	1-10

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search
04 September, 2003 (04.09.03)

Date of mailing of the international search report
16 September, 2003 (16.09.03)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/07923

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 10-274640 A (Hitachi, Ltd.), 13 October, 1998 (13.10.98), Full text; Figs. 1 to 10 & US 6008490 A	1-10
A	SPAHR, C.S. et al., Simplification of complex peptide mixtures for proteomic analysis: Reversible biotinylation of cysteinyl peptides. Electrophoresis., May 2000, Vol.21, No.9, pages 1635 to 1650	1-10

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl. 7 G01N27/62, H01J49/26

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl. 7 G01N27/62-27/70, H01J49/00-49/48,
G01N33/48-33/98, C12Q1/00-1/70

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2003年
日本国登録実用新案公報	1994-2003年
日本国実用新案登録公報	1996-2003年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	WO 95/25281 A1 (UNIVERSITY OF WASHINGTON) 1995. 09. 21, 全文, 第1-6図 & JP 9-510780 A	1-10
A	JP 8-124519 A (株式会社島津製作所) 1996. 05. 17, 全文, 第1-3図 (ファミリーなし)	1-10
A	JP 5-164751 A (株式会社島津製作所) 1993. 06. 29, 全文, 第1-3図 (ファミリーなし)	1-10

☒ C欄の続きにも文献が列举されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」 口頭による開示、使用、展示等に言及する文献
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」 同一パテントファミリー文献

国際調査を完了した日

04. 09. 03

国際調査報告の発送日

16.09.03

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)
郵便番号100-8915
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

鈴木 俊光



2W

9115

電話番号 03-3581-1101、内線 3292

C (続き) 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 10-274640 A (株式会社日立製作所) 1998. 10. 13, 全文, 第1-10図 & US 6008490 A	1-10
A	SPAHR, C.S. et. al. Simplification of complex peptide mixtures for proteomic analysis: Reversible biotinylation of cysteinyl peptides. Electrophoresis. May 2000, Vol. 21, No. 9, pages 1635 - 1650	1-10